Discrete Mathematics 23 (1978) 67–71.

NOTE

MAXIMAL FREQUENCIES OF EQUIVALENCES WITH SMALL CLASSES

Jiří DEMEL

Katedra Ekonomiky a Řízení Stavebnicktví, Stavební Fakulta ČVUT, Zikova 4, 166 29 Praha 6, Czechoslovakia

Received 13 December 1976

If e, E are equivalences on a finite set, then the frequency of e in E is the number of subequivalences of E which are isomorphic to e. Given an equivalence e on a certain set the problem arises of finding an equivalence E such that the frequency of e in E is maximal. In this paper we solve this problem for those equivalences e all of whose classes have cardinality at most 2.

In the whole paper every set is assumed to be finite. Denote by $\mathscr{E}_M$ the set of equivalences on a set M. For a given equivalence e define a function $F_M(e, -)$: $\mathscr{E}_M \to \mathcal{N}$ by $F_M(e, E) = \operatorname{card}\{U \subset M : E \restriction U \text{ is isomorphic to } e\}$ ($E \restriction U$ means the partialization of E on U). We shall call the number $F_M(e, E)$ the *frequency* of e in E.

When solving some statistical problems by the method of maximal probability the question of characterizing equivalences in which the frequency of an equivalence e is maximal, can occur. Our aim is to solve this problem for equivalences containing only small classes, i.e., classes the cardinality of which is not more than two.

We shall prove that if the frequency of e in E is maximal, e having small classes, then E has "almost equal" classes. Moreover, for "big" sets M an upper and lower bound of the number of non-empty classes of E will be given.

Denote by $e_{k,p}$ an arbitrary equivalence on $2k+p$ elements which has k classes with cardinality 2 and p classes with cardinality 1.

We shall say that an equivalence E has *almost equal classes* if for any two non-empty classes, V, W one has $|\operatorname{card}(V) - \operatorname{card}(W)| \leq 1$.

Theorem 1. *If $2k+p>1$ and the frequency $F_M(e_{k,p}, -)$ is maximal in E, then E has almost equal classes.*

Remark 2. If $k=0$ and $p=1$ then $F_M(e_{k,p}, E) = \operatorname{card}(M)$ for any equivalence E on M.

Remark 3. Theorem 1 does not hold for equivalences e containing classes with cardinality greater than 2. For example, let e have one class with cardinality 1 and one class with cardinality 3. Then for card $(M)=9$ the frequency is maximal in an equivalence which has one class with cardinality 2 and one class with cardinality 7.

Theorem 4. *For arbitrary natural numbers k, p, where $k>0$, there exists a natural number N_0 such that for every set M with* card $(M)>N_0$ *one has*

$$\frac{p(p-1)}{2k}+\tfrac{3}{2}p+k-1\leq m\leq\frac{p(p-1)}{2k}+2p+2k-1$$

where m is the number of non-empty classes of an equivalence E on M, in which the frequency of $e_{k,p}$ is maximal.

Remark 5. If $k=0$ and $p>1$, then $m=\operatorname{card}(M)$. If $k=0$ and $p=1$, then $1\leq m\leq\operatorname{card}(M)$. It is surprising that for $k>0$ and card $(M)>N_0$ the number m depends only on k and p and does not depend on card (M).

To prove Theorem 1 and Theorem 4 let us introduce some notions and notations.

Let E be an equivalence on M; take two classes of E such that at least one is non-empty. Denote by A the union of these classes; $B=M\backslash A$.

If $i+1<j$ we shall denote by $\mathcal{U}_A(E)$ an arbitrary equivalence on M such that $\mathcal{U}_A(E)\restriction B=E\restriction B$ and the set A contains exactly two classes of $\mathcal{U}_A(E)$, the cardinalities of which are $j-1$ and $i+1$.

Similarly, if $i\neq 0$, we shall denote by $\mathcal{D}_A(E)$ an arbitrary equivalence on M such that $\mathcal{D}_A(E)\restriction B=E\restriction B$ and the set A contains two classes of $\mathcal{D}_A(E)$, the cardinalities of which are $j+1$ and $i-1$. It may happen that one of these classes is empty.

Denote $\mathcal{U}_A^{r+1}(E)=\mathcal{U}_A(\mathcal{U}_A^r(E))$, if the right side has a sense. Similarly for $\mathcal{D}_A^{r+1}(E)$.

For simplicity we shall write $G(k,p)$ instead of $F_B(e_{k,p},E\restriction B)$. We then have:

$$\begin{aligned}F_M(e_{k,p},E)&=ijG(k,p-2)+(i+j)G(k,p-1)+G(k,p)\\&\quad+\left(\binom{i}{2}+\binom{j}{2}\right)G(k-1,p)+\left(\binom{i}{2}j+\binom{j}{2}i\right)G(k-1,p-1)\\&\quad+\binom{i}{2}\binom{j}{2}G(k-2,p).\end{aligned}$$

The expressions for the frequencies $F_M(e_{k,p},\mathcal{U}_A(E))$, resp. $F_M(e_{k,p},\mathcal{D}_A(E))$ are similar except that i, j must be replaced by $i+1$, $j-1$, resp. $i-1$, $j+1$.

If the equivalence $\mathcal{U}_A(E)$ exists, put

$$\begin{aligned}Y^{\mathcal{U}}_{E,A}(i,j)&=F_M(e_{k,p},\mathcal{U}_A(E))-F_M(e_{k,p},E)\\&=(j-i-1)[G(k,p-2)-G(k-1,p)\\&\quad+(\tfrac{1}{2}(i+j)-1)G(k-1,p-1)+\tfrac{1}{2}i(j-1)G(k-2,p)].\end{aligned}$$

Designate by $Z^{\mathscr{U}}_{E,A}(i,j)$ the expression in brackets. Similarly, if the equivalence $\mathscr{D}_A(E)$ exists, put

$$\begin{aligned}Y^{\mathscr{D}}_{E,A}(i,j) &= F_M(e_{k,p}, \mathscr{D}_A(E)) - F_M(e_{k,p}, E)\\ &= (j-i-1)[G(k,p-2) - G(k-1,p)\\ &\quad + (\tfrac{1}{2}(i+j)-1)G(k-1,p-1) + \tfrac{1}{2}j(i-1)G(k-2,p)].\end{aligned}$$

Designate by $Z^{\mathscr{D}}_{E,A}(i,j)$ the expression in brackets.

Since $i \leqslant j$, we have $i(j-1) \geqslant j(i-1)$, therefore $Z^{\mathscr{U}}_{E,A}(i,j) \geqslant Z^{\mathscr{D}}_{E,A}(i,j)$.

Lemma 6. *If $Z^{\mathscr{U}}_{E,A}(i,j) > 0$, then $F_M(e_{k,p}, E) < F_M(e_{k,p}, \mathscr{U}^r_A(E))$, where $r > 0$ is a natural number $\leqslant \frac{1}{2}(j-i)$.*

If $Z^{\mathscr{D}}_{E,A}(i,j) < 0$, then $F_M(e_{k,p}, E) < F_M(e_{k,p}, \mathscr{D}^r_A(E))$, where $1 \leqslant r \leqslant i$.

Proof. Let $Z^{\mathscr{U}}_{E,A}(i,j) > 0$. One has $j-i-1 > 0$, otherwise the expression $Z^{\mathscr{U}}_{E,A}(i,j)$ is not defined. Therefore $Y^{\mathscr{U}}_{E,A}(i,j) > 0$, thus $F_M(e_{k,p}, E) < F_M(e_{k,p}, \mathscr{U}_A(E))$. The expressions $Z^{\mathscr{U}}_{\mathscr{U}_{A'}(E),A}(i,j)$ are defined for $r \leqslant \frac{1}{2}(j-i)$ and $Z^{\mathscr{U}}_{\mathscr{U}_{A'}(E),A}(i,j) = Z^{\mathscr{U}}_{E,A}(i+v, j-v) \geqslant Z^{\mathscr{U}}_{E,A}(i,j) > 0$. Hence $F_M(e_{k,p}, \mathscr{U}^{r-1}_A(E)) < F_M(e_{k,p}, \mathscr{U}^r_A(E))$. Thus the proof of the first part of Lemma 6 is finished.

The proof of the second part of Lemma 6 is similar.

Proof of Theorem 1. Assume a contrary. Let E contain two non-empty classes V, W such that $\operatorname{card}(W) - \operatorname{card}(V) \geqslant 2$. We shall denote $i = \operatorname{card}(V)$, $j = \operatorname{card}(W)$, $A = W \cup V$. Then the equivalences $\mathscr{U}_A(E)$ and $\mathscr{D}_A(E)$ exist and the expressions $Z^{\mathscr{U}}_{E,A}(i,j)$, $Z^{\mathscr{D}}_{E,A}(i,j)$ are defined.

If $Z^{\mathscr{U}}_{E,A}(i,j) > 0$, then by Lemma 6 the frequency of $e_{k,p}$ in $\mathscr{U}_A(E)$ is greater than the frequency in E.

If $Z^{\mathscr{D}}_{E,A}(i,j) < 0$, then the frequency of $e_{k,p}$ in $\mathscr{D}_A(E)$ is greater than the frequency in E.

If $Z^{\mathscr{U}}_{E,A}(i,j) = Z^{\mathscr{D}}_{E,A}(i,j) = 0$, then $i = j$, a contradiction with $j - i \geqslant 2$.

Proof of Theorem 4. By Theorem 1, E has almost equal classes. Among all equivalences on M which have this property, there may exist equivalences the frequency in which is smaller than in those constructed from them by using $\mathscr{U}$ or $\mathscr{D}$. We get the upper bound by using $\mathscr{D}$ and the lower bound by using $\mathscr{U}$.

At first we shall investigate cases with $k+p<3$. The assertion is obvious for $k=1$ and $p=0$, because $m=1$.

If $k=p=1$, then E has exactly two non-empty classes. Assuming E has at least three non-empty classes, we can choose the set $A \subset M$ such that $i \leqslant j \leqslant \operatorname{card}(B)$. Then $Z^{\mathscr{D}}_{E,A}(i,j) = \frac{1}{2}(i+j) - 1 - \operatorname{card}(B) < 0$, a contradiction to Lemma 6.

Let $k=2$ and $p=0$. Then E has two or three non-empty classes. Assuming E has at least four non-empty classes, we can choose the set A in such a way that B contains the greatest class of E. Denote q the cardinality of this class. Then $G(1,0) > \binom{q}{2} \geqslant \binom{j}{2} \geqslant \frac{1}{2}j(i-1)$. Hence $Z^{\mathscr{D}}_{E,A}(i,j) = \frac{1}{2}j(i-1) - G(1,0) < 0$, a contradiction to Lemma 6.

Now, let $k+p\geq 3$. At first we shall prove the upper bound. The equivalence E has at least three non-empty classes, hence we can choose the set A such that it contains exactly two non-empty classes of equal size. Denote by n the number of non-empty classes contained in B; suppose that r of these classes have cardinality $a-1$ and $n-r$ classes have cardinality a. We have $i=j$ and either $i=a$ or $i=a-1$.

Then the following holds:

$$G(k,p-2)\leq\binom{n}{k}\binom{a}{2}^k\binom{n-k}{p-2}a^{p-2},$$

$$G(k-1,p)\geq\binom{n}{k-1}\binom{a-1}{2}^k\binom{n-k+1}{p}(a-1)^p,$$

$$G(k-1,p-1)\leq\binom{n}{k-1}\binom{a}{2}^{k-1}\binom{n-k+1}{p-1}a^{p-1},$$

$$G(k-2,p)\leq\binom{n}{k-2}\binom{a}{2}^{k-2}\binom{n-k+2}{p}a^p.$$

Then

$$Z^{\mathscr{D}}_{E,A}(i,j)\leq\frac{n!}{k!\,p!\,(n-k-p+2)!}\Bigg[\binom{a}{2}^{k-1}a^{p-1}(\tfrac{1}{2}p(p-1)(a-1)$$
$$+kp(a-1)+ak(k-1))-\binom{a-1}{2}^{k-1}(a-1)^{p-1}k(n-k-p+2)\Bigg].$$

Denote $N=\operatorname{card}(M)$. One has $(a-1)(n+2)\leq N\leq a(n+2)$, so that $N/(n+2)\leq a\leq(N+n+2)/(n+2)$, hence $a=(N+b)/(n+2)$, where $0\leq b\leq n+2$. After substitution we get

$$Z^{\mathscr{D}}_{E,A}(i,j)\leq\frac{n!}{k!\,p!\,(n-k-p+2)!}\Bigg[\frac{(N+b)^{k+p-2}(N+b-n-2)^{k-1}}{2^{k-1}(n+2)^{2k+p-2}}$$
$$\cdot\left((N+b-n-2)\left(\binom{p}{2}+kp\right)+(N+b)k(k-1)\right)$$
$$-\frac{(N+b-n-2)^{k+p-2}(N+b-2n-4)^{k-1}}{2^{k-1}(n+2)^{2k+p-2}}k(n-k-p+2)\Bigg].$$

The expression on the right hand side of the inequality is a polynomial in N of degree $2k+p-2$, the leading coefficient of which is $\frac{1}{2}p(p-1)+kp+k(k-1)-k(n-k-p+2)$.

If this coefficient is negative, then for N great enough we have $Z^{\mathscr{D}}_{E,A}(i,j)<0$, a contradiction to Lemma 6. Therefore the coefficient is non-negative, hence

$$\frac{p(p-1)}{2k}+2k+2p-1\geq n+2=m$$

and the proof of the upper bound is complete.

The proof of the lower bound is similar. Choose A such that it contains exactly one non-empty class of E. We have $i=0$ and either $j=a$ or $j=a-1$.

One has

$$G(k,p-2)\geqslant\binom{n}{k}\binom{a-1}{2}^k\binom{n-k}{p-2}(a-1)^{p-2},$$

$$G(k-1,p)\leqslant\binom{n}{k-1}\binom{a}{2}^{k-1}\binom{n-k+1}{p}a^p,$$

$$G(k-1,p-1)\geqslant\binom{n}{k-1}\binom{a-1}{2}^{k-1}\binom{n-k+1}{p-1}(a-1)^{p-1},$$

$$G(k-2,p)\geqslant\binom{n}{k-2}\binom{a-1}{2}^{k-2}\binom{n-k+2}{p}(a-1)^p.$$

Thus

$$Z^{\mathcal{U}}_{E,A}(i,j)\geqslant\frac{n!}{k!\,p!\,(n-k-p+2)!}\cdot\left[\binom{a-1}{2}^{k-1}(a-1)^{p-1}\left(\binom{p}{2}(a-2)+kp(\tfrac{1}{2}a-1)\right)-\binom{a}{2}^{k-1}a^p k(n-k-p+2)\right].$$

The equivalence E has now $n+1$ classes, then $(a-1)(n+1)\leqslant N\leqslant a(n-1)$, hence $a=(N+b)/(n+1)$, where $0\leqslant b\leqslant n+1$. After substitution we get:

$$Z^{\mathcal{U}}_{E,A}(i,j)\geqslant\frac{n!}{k!\,p!\,(n-k-p+2)!}\left[\frac{(N+b-n-1)^{k+p-2}(N+b-2n-2)^{k-1}}{2^{k-1}(n+1)^{2k+p-2}}\cdot\left(\binom{p}{2}(N+b-2n-2)+kp(\tfrac{1}{2}(N+b)-1)\right)-\frac{(N+b)^{k+p-1}(N+b-n-1)^{k-1}}{2^{k-1}(n+1)^{2k+p-2}}k(n-k-p+2)\right].$$

The expression on the right hand side of the inequality is a polynomial in N of degree $2k+p-2$, the leading coefficient of which is $\frac{1}{2}p(p-1)+\frac{1}{2}kp-k(n-k-p+2)$.

If this coefficient is positive, then for N great enough we have $Z^{\mathcal{U}}_{E,A}(i,j)>0$, a contradiction to Lemma 6. Therefore the coefficient is non-positive, hence

$$\frac{p(p-1)}{2k}+\frac{3}{2}p+k-1\leqslant n+1=m.$$

Thus the proof is finished.